第四章

随机变量的数学特征

第四章 随机变量的数字特征

4.1 数学期望

1、离散型的期望

$E_X=\sum_{k=1}^{\infty} x_k p_k$　　级数 $\sum_{k=1}^{\infty} x_k p_k$ 绝对收敛

2、连续型的期望

$\int_{-\infty}^{+\infty} x f(x)dx$ 绝对收敛，则 $E_X=\int_{-\infty}^{+\infty} x f(x)dx$　　x：取值　$f(x)$：概率

- 0-1分布 $E_X=p$　　二项分布 $X\sim b(n,p)$　$E(X)=np$
- 泊松分布 $X\sim\pi(\lambda)$　$E(X)=\lambda$　　$X\sim U(a,b)$ 平均分布 $E(X)=\frac{a+b}{2}$
- 参数为 θ 的指数分布 $E(X)=\theta$　　正态分布 $X\sim N(\mu,\sigma^2)$　$E(X)=\mu$

3、随机变量函数的期望

X. $Y=g(X)$. $E_X=\sum x_i p_i$　　$E_Y=\sum g(x_i)p_i$

$\int_{-\infty}^{+\infty} x f(x)dx$　　$\int_{-\infty}^{+\infty} g(x)f(x)dx$

二维：

(x,y)　$Z=g(X,Y)$　$E_Z=\sum_i\sum_j g(x_i,y_j)p_{ij}$

$E_Z=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f(x,y)dxdy$

$f(x,y)$.

$E(Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y f(x,y)dydx$

$=\int_{-\infty}^{+\infty} y f_Y(y)dy$

$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx$

4、数学期望的性质

1、$E_C=C$（C为常数）　　2、$E(X+C)=EX+C$

3、$E(CX)=CE_X$　　4、$E(kx+b)=kEx+c$

5、$E(X\pm Y)=EX\pm EY$　　6、独立时 $E(XY)=EX\cdot EY$

（$E(\sum C_iX_i)=\sum C_iEX_i$）

（$E(\frac{1}{n}\sum X_i)=\frac{1}{n}\sum EX_i$）

5、条件期望：一变量取某值，另一变量期望

(1) $E(X|Y=y_i)=\sum x_iP(X=x_i|Y=y_i)$

$E(Y|X=x_i)=\sum y_iP(Y=y_j|X=x_i)$

X \ Y	1	2	3
0	0.1	0.2	0.3
1	0.2	0.1	0.1

$E(Y|X=1)$

Y	1	2	3
P	0.5	0.25	0.25

$=1\times0.5+2\times0.25+3\times0.25=1.75$

(2) 连续 $E(X|Y=y)=\int_{-\infty}^{+\infty} x f(x|y)dx$

$E(Y|X=x)=\int_{-\infty}^{+\infty} y f(y|x)dy$

4.2 方差

1、定义　$D(X)=E(X-EX)^2$ 方差　$\sqrt{DX}$ 标准差

离散 $D(X)=\sum_k(X_k-EX)^2P_k$　连续 $D(X)=\int_{-\infty}^{+\infty}(x-EX)^2f(x)dx$

$D(X)=E(X^2)-E^2(X)$

2、性质

- (1) $D(C)=0$　(2) $D(X+C)=D(X)$　(3) $D(CX)=C^2D(X)$
- (4) $D(kx+b)=k^2D(X)$　(5) $X\cdot Y$独立 $D(X\pm Y)=D(X)+D(Y)$
- (6) $D(X)=0\Leftrightarrow P(X=E(X))=1$　(7) 独立时 $D(X_1+X_2\cdots)=DX_1+\cdots+\cdots$

- (1) $E(C)=C$　(2) $E(X+C)=E(X)+C$　(3) $E(CX)=CE(X)$
- (4) $E(kx+b)=kE(x)+b$　(5) $E(X\pm Y)=E(X)\pm E(Y)$
- (6) $E(X\cdot Y)=EX\cdot EY$（独立时）

$X^*=\frac{X-E(X)}{\sqrt{DX}}$　$E(X^*)=0$　$D(X^*)=1$

3、离散型

0-1分布　$E(X)=p$　$D(X)=p(1-p)$

二项分布　$E(X)=np$　$D(X)=np(1-p)$

几何分布　$E(X)=\frac{1}{p}$　$D(X)=\frac{1-p}{p^2}$

泊松分布　$E(X)=\lambda$　$D(X)=\lambda$

4、连续型

均匀分布　$E(X)=\frac{a+b}{2}$　$D(X)=\frac{(b-a)^2}{12}$

指数分布　$E(X)=\frac{1}{\lambda}$　$D(X)=\frac{1}{\lambda^2}$

正态分布　$E(X)=\mu$　$D(X)=\sigma^2$　$X\sim N(\mu,\sigma^2)$

切比雪夫不等式　$P\{|X-E(X)|\geq\varepsilon\}\leq\frac{\sigma^2}{\varepsilon^2}$

σ^2越小，$\{|X-E(X)|<\varepsilon\}$概率越大，

即 X 集中在期望附近可能性越大

4.4 协方差

$cov(X,Y)=E[(X-E(X))(Y-E(Y))]=E(XY)-E(X)E(Y)$

$D(X\pm Y)=D(X)+D(Y)\pm2cov(X,Y)$

性质　$cov(X,Y)=cov(Y,X)$

$cov(aX,bY)=ab\,cov(X,Y)$

$cov(X_1+X_2,Y)=cov(X_1,Y)+cov(X_2,Y)$

$cov(C,X)=0$　　$cov(X,Y)=0$（$X\cdot Y$独立）

$cov(X,Y)$受单位影响　$X^*=\frac{X-E(X)}{\sqrt{DX}}$　$Y^*=\frac{Y-E(Y)}{\sqrt{DY}}$ 标准化

可消除影响

$cov(X^*,Y^*)=\frac{cov(X,Y)}{\sqrt{DX}\sqrt{DY}}=\rho$（不受单位影响）

4.4 相关系数

$\rho=\frac{cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=\frac{E(XY)-E(X)E(Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$ 相关系数，与协方差同号

$|\rho|\leq1$　引理：$[E(XY)]^2\leq E(X^2)\cdot E(Y^2)$，可证

$|\rho|=1\Leftrightarrow X$与$Y$以$P=1$存线性关系，$P(Y=aX+b)=1$

ρ用于衡量XY之间的线性关系

① $\rho=1$，XY完全正相关

② $\rho=-1$，XY完全负相关

③ $|\rho|\to0$，XY线性关系弱

④ $\rho=0$，XY不存在线性关系

XY不相关（线性不相关）　　XY独立（无任何关系，包括线性/非线性）

① XY独立，则XY不相关　② XY不相关，XY不一定独立

4.5 中心矩与原点矩

原点矩 $E(X^k)$　期望$E(X)$，一阶原点矩

中心矩 $E(X-EX)^k$　以$E(X)$为中心.

二阶中心矩为方差　一阶中心矩$=0$

原点矩

- ① 离散 $\sum x_i^k p_i$
- ② 连续 $\int_{-\infty}^{+\infty}x^kf(x)dx$

中心矩

- ① 离散 $\sum(x_i-E(x))^kp_i$
- ② 连续 $\int_{-\infty}^{+\infty}(x-E(x))^kf(x)dx$

$E(X^kY^L)$，是X和Y的$k+L$阶混合矩

$E\{[X-E(X)]^k[Y-E(Y)]^L\}$，$k+L$阶混合中心矩

协方差矩阵 $\begin{pmatrix}C_{11}&C_{12}\\C_{21}&C_{22}\end{pmatrix}$　$C_{11}=E\{[X_1-E(X_1)]^2\}$

$C_{12}=E\{[X_1-E(X_1)][X_2-E(X_2)]\}$

本笔记在https://github.com/dydcyy-gh/study-notes开源